Multi-scale similarity entropy as a complexity descriptor to discriminate healthy to distress foetus

Jean Marc Girault, Souad Oudjemia, Iulian Voicu

Multi-Scale Similarity Entropy as a Complexity Descriptor to discriminate Healthy to Distress fetus

Jean-Marc Girault[a,], Souad Oudjemia[a], Iulian Voicu[a]

[a] *Université François Rabelais de Tours, UMR-S930, Tours, France*

Abstract

This paper deals with the discrimination between suffering foetuses and normal foetuses by means of a multi-scale similarity entropy. Sample entropy and similarity entropy are evaluated in multi-scale analysis on foetal heart rate signals. Without multi-scale analysis, our results show that only the similarity entropy differentiate suffering foetuses to normal foetuses. Furthermore with the multi-scale analysis, our results show that both the sample entropy and the similarity entropy can discriminate the distressed foetuses to normal foetuses. In all cases the similarity entropy outperforms the sample entropy that is encouraging for another biomedical applications.

Keywords : Entropy, Similarity, complexity, time series.

1. INTRODUCTION

One of the most important purpose in foetal monitoring antepartum is the detection of the foetal distress since it can help the obstetrician to rapidly decide whether it is necessary or not to deliver the foetus by caesarean section. However this is not the only one reason that explains its usefulness since the obstetrician needs also to understand the fundamental reasons leading to such a foetal distress. This quest was made difficult due to the wide range of possible pathologies and the different degrees of severity and symptoms encountered during the gestational period. Although the foetal distress was in a general way characterized by a reduction of the maternal-foetal respiratory exchanges leading the foetus metabolism to an anaerobic state, all the underlying reasons were not yet elucidated.

Email address: jean-marc.girault@univ-tours.fr (Jean-Marc Girault)

Recently, it has been shown by [1] that a foetal heart rate (FHR) monitoring completed by a certain numbers of indicators could reduce the morbidity and mortality rate up to 7 per thousand. Among all these proposed indicators, the most simple and commonly used were those based on the estimation of the FHR acceleration-deceleration and based on the short and mean terms FHR variability to name a few. The second class of indicators was based on the assessment of the FHR complexity. The most well-known descriptors were based on different kinds of entropy as for instance the approximate entropy or the sample entropy (SamEnt). These descriptors developed for characterizing time series without any *a priori* information on the nature of the underlying system, measured the degree of regularity of the analysed time series. These measures were realized by evaluating the probability of finding similar m-patterns that are templates composed of m samples.

Recently, the work performed by [2] showed that the multi-scale analysis could still improve the characterization of physiological time series. This multi-scale analysis of entropy was then applied by [3], [4] for studying the suffering foetuses. This work that was the starting point of our study showed that it was possible to discriminate suffering foetuses to healthy foetuses.

Our study aimed to improve the differentiation between normal and suffering foetuses by investigating deeper the links between m-patterns. To do so we investigated the similarity entropy previously introduced by [5] and we extended its definition with a multi-scale point view. By assessing the multi-scale similarity entropy, we showed that it was possible to better discriminate healthy foetuses to suffering foetuses in comparison to the multi-scale sample entropy.

To prove it, we firstly present the materials used to measure the foetal heart rate (FHR), secondly we present the multi-scale similarity analysis and thirdly we present some results derived from our dataset. We close this research work by a discussion and a conclusion.

2. MATERIALS

Our system was composed of a personal computer and our home-made Actifoetus unit. The Actifoetus unit contained three groups of four transducers and a Doppler acquisition board. The transducers exploring the foetal heart were non-focused and mono-element. They were circular in shape, with a diameter of 13.5 mm and an acoustic power of 1 mW/cm^2. Geometrically, the transducers were located at the center of gravity and at the tops of an equilateral triangle of sides measuring 40.7 mm.

The transducers were placed on the mother's abdomen. They transmitted a sinusoidal pulse at 2.25 MHz with a pulse repetition frequency of 1kHz. The wave was propagated through the mother's abdomen towards the foetal heart. The backscattered signal was recorded from five different depths, annotated D1... D5. Note, only one channel was considered in the present study.

The ultrasound signal received was converted into an electrical signal and amplified to compensate the attenuation of 1 dB/cm/MHz. The signal was then

Figure 1: Illustration of the fundamental difference between sample and similarity entropies. Top of the figure, identification of the first 2-pattern into the time series for the sample entropy process. Bottom of the figure, identification of the first centred 2-pattern into the time series for the similarity entropy process.

demodulated in phase (I) and quadrature (Q). After demodulation, the signals were digitized. The digital outputs of the converters represented the digital Doppler signal.

The Doppler signals were acquired at CHRU "Bretonneau" Tours, France. The consent of each patient was obtained and the study was approved by the ethics committee of the Clinical Investigation Centre for Innovative Technology of Tours (CIC-IT 806 CHRU of Tours). Patients are older than eighteen years and pregnancy was singular. Six patients (three normal foetuses and three suffering foetuses) were included in this study. Foetuses whose gestational age ranged from the 25th to the 39th week were monitored during 30 minutes. FHR was evaluated as proposed by [6] that is every 250 ms leading to 7200 samples for a recording of 30 minutes.

3. METHODS

From FHR recordings, each time series $x(n)$ composed of $M = 7200$ points were represented by a vector $\mathbf{X}$. To catch the fluctuations present in the time series at different scales, the multi-scale analysis was introduced. This method consisted in evaluating approximate versions of the original time series from a local average of neighbouring points. This procedure proposed by [2] is named "coarse-grained". The new time series composed of M/α samples at the scale α writes:

$$y_\alpha(k) = \sum_{i=(k-1)\alpha+1}^{k\alpha} x_i, \tag{1}$$

for $1 \leq k \leq M/\alpha$, $y_1(k) = x(k)$ being the original time series. To ensure the same length for each time series after the coarse-grained process, time series were interpolated by a factor α. The vector of the interpolated times series, composed of M samples, was then obtained. For each vector a short term similarity measure was performed on a sub-vector $\mathbf{Y}_\alpha$ composed of $N = 720$ points (3 min.).

The method that was tested here was based on a non multi-scale analysis initially proposed by [5]. Unlike conventional approaches where the probability of finding a m-pattern whose amplitude was similar, in this method the probability of finding similar centred m-patterns was evaluated instead. This centring process that seems to the untrained eye looks completely benign enabled to compare similar m-patterns located everywhere into the time series while for conventional approaches only m-patterns located at the same amplitude were accounted. The originality of our proposed method was to extend the method

proposed by [5] with a mutli-scale point of view. Note that a preliminary study on this subject was presented by [7].

In this paper two kinds of entropy were evaluated: the sample entropy (SamEnt) and the sample entropy of centred m-patterns that was named similarity entropy (SimEnt). This pattern centring process came down to remove the FHR base-line of the FHR, it was similar to a detrending process.

To exemplify, an illustration was proposed in Fig.1. At the top of Fig.1 a part of the sample entropy process was presented. In this case, the probability of finding the first 2-pattern was evaluated and only two 2-patterns were found similar. At the bottom of Fig.1 a part of the similarity process was presented. In this case the probability of finding the first centred 2-pattern was evaluated and eight centred 2-patterns were found similar. The similarity seeking process was thus a matter of studying fluctuations around the trend.

From the time series vector $\mathbf{Y}_\alpha$, a vector sequence was formed:

$$\mathbf{Y}_{\alpha,i}^{(m)} = \{y_\alpha(i), y_\alpha(i+1), ..., y_\alpha(i+m-1)\} - \bar{\mathbf{Y}}_{\alpha,i}^{(m)} \tag{2}$$

with $\bar{\mathbf{Y}}_{\alpha,i}^{(m)} = \frac{1}{m}\sum_{l=0}^{m-1} y_\alpha(i+l)$ for i ranging from 1 to $N-m+1$. $\mathbf{Y}_{\alpha,i}^{(m)}$ represents m consecutive y_α values starting from the ith point, m being the length of the pattern. The distance $d_{i,j}^{(m)}$ between $\mathbf{Y}_{\alpha,i}^{(m)}$ and $\mathbf{Y}_{\alpha,j}^{(m)}$ is defined as:

$$d_{i,j}^{(m)} = d\left[\mathbf{Y}_{\alpha,i}^{(m)}, \mathbf{Y}_{\alpha,j}^{(m)}\right] = \max_{k\in(0,m-1)} \left|y_\alpha(i+k) - y_\alpha(j+k)\right|. \tag{3}$$

The average degree of similarity between $\mathbf{Y}_{\alpha,i}^{(m)}$ and its neighbouring vectors $\mathbf{Y}_{\alpha,j}^{(m)}$ within tolerance r can then be defined as

$$B_i^{(m)}(r) = \frac{1}{N-m-1}\sum_{j=1,j\neq i}^{N-m} \Theta(d_{i,j}^{(m)} - r) \tag{4}$$

where Θ is the Heaviside function, j ranges from 1 to $(N-m)$ and $j \neq i$ to exclude self-matches.

When vector $\mathbf{Y}_{\alpha,i}^{(m)}$ is very close to its neighbouring vector $\mathbf{Y}_{\alpha,j}^{(m)}$, that is when the distance $d_{i,j}^{(m)} < r$, then a pattern match occurs. Finally, define the function $B^{(m)}(r)$ as follows:

$$B^{(m)}(r) = \frac{1}{N-m}\sum_{i=1}^{N-m} B_i^{(m)}(r). \tag{5}$$

Similarly, form the vector sequence $\mathbf{Y}_{\alpha,i}^{(m+1)}$ and get the function $A^{(m)}(r)$

$$A^{(m)}(r) = \frac{1}{N-m}\sum_{i=1}^{N-m} A_i^{(m)}(r), \tag{6}$$

Figure 2: Two foetuses heart rate derived from our dataset. Top of the figure, foetal heart rate of a normal foetus. Bottom of the figure, foetal heart rate of a distressed foetus.

where

$$A_i^{(m)}(r) = \frac{1}{N-m-1} \sum_{j=1, j\neq i}^{N-m} \Theta(d_{i,j}^{(m)} - r). \tag{7}$$

The similarity entropy $SimEnt(m, r, N)$ for a finite dataset of N samples writes at a scale α:

$$SimEnt(\alpha, m, r, N) = log(B^{(m)}) - log(A^{(m)}). \tag{8}$$

Note that the similarity entropy possesses the same definition of the sample entropy, the one difference was on the definition of the m-patterns which were centred for the similarity entropy.

4. APPLICATION TO THE FOETAL HEART RATE SIGNAL

From our dataset composed on six recordings of 30 minutes, three recordings corresponded to healthy foetuses and three to distressed foetuses. As an illustration, it was reported on Fig.2 two FHRs. On the top of Fig.2 was reported the FHR of a normal foetus while in the bottom of Fig.2 was depicted a FHR of a suffering foetus. In the both cases, the FHR fluctuated around 135 beats per minutes. Note that the FHR of the suffering foetus seemed to be more variable. For convenient, each FHR recordings were centred and normalized. For each recordings, a short term analysis was performed on the vector $\mathbf{Y}_{\alpha,i}^{(m)}$ by using a sliding window composed of 720 points with an overlap of 97 %. $SamEnt(\alpha, m, r, N)$ and $SimEnt(\alpha, m, r, N)$ were evaluated for each extracted signal at different scale α ranging from 1 to 6 and for $m = 2$, $r = 0.1$, $N = 720$. For the healthy group and for the suffering group, each estimation of sample entropy and similarity entropy were gathered in a view to estimate the mean and the standard deviation of each estimation. In Fig.3 was reported such estimations. On the top of Fig.3, the sample entropy for normal and suffering foetuses was reported in regards with the scale. On the bottom of Fig.3, the similarity entropy for normal and suffering foetuses was reported in regards with the scale. In both cases, the sample entropy and the similarity entropy decreased as the scale increased. This behaviour seemed corroborate the fact that as the scale increased, the time series became smoother thus reducing its complexity. However this behaviour was not confirmed by the work of [3, 4] where the sample entropy increased with the growing of the scale.

In any event, note that it was impossible to discriminate suffering to normal foetuses by comparing the sample entropy for each scale value independently whereas it was the case for the similarity entropy with a scale ranging from 1 to 2.

Figure 3: Multi-scale analysis of the sample and similarity entropies (2-pattern) for normal and suffering foetuses. Top of the figure, sample entropy versus scale for normal foetuses (dot line) and suffering foetuses (continuous line). Bottom of the figure, similarity entropy versus scale for normal foetuses (dot line) and suffering foetuses (continuous line).

Finally, note that both the sample entropy and the similarity entropy were higher for distressed foetuses than for healthy foetuses. This point seemed to be corroborated by the more fluctuating nature of FHR for suffering foetuses than for normal foetuses as obtained in Fig.2.

With this in mind, we can now turn to fully take into account the multi-scale analysis. As proposed by [4], it seemed much more interesting to evaluate the deviation of both descriptors at different scale. To do so we evaluated the relative error between similarity entropies and the sample entropies at different scales α and α':

$$RE_1(\alpha, \alpha') = \frac{\overline{SimEnt(\alpha)} - \overline{SimEnt(\alpha')}}{\overline{SimEnt(\alpha)}}, \tag{9}$$

$$RE_2(\alpha, \alpha') = \frac{\overline{SamEnt(\alpha)} - \overline{SamEnt(\alpha')}}{\overline{SamEnt(\alpha)}}, \tag{10}$$

where $\overline{SimEnt(\alpha)}$ and $\overline{SamEnt(\alpha)}$ were the mean values at each scale.

In table 1, $RE_1(1,6)$ and $RE_2(1,6)$ were reported. Findings derived from table 1, showed that it was possible to separate healthy to suffering foetus by using $\alpha = 1$,$\alpha' = 6$. In such a case, a deviation of $56\% - 41\% = 15\%$ and a deviation of $74\% - 51\% = 23\%$ were obtained, respectively. By using the multi-scale analysis, it was easier to differentiate with similarity entropy than with sample entropy. This result confirmed that the slope of the similarity entropy obtained for the normal and the suffering foetus were very different (see the bottom of Fig.3) to enable the discrimination.

5. DISCUSSION AND CONCLUSION

The multi-scale analysis of similarity or sample entropies, has shown its efficiency to discriminate suffering to normal foetuses. In this study this was the multi-scale similarity entropy that gave the best performances and thus seemed to be more adapted to our dataset. Even if our findings did not match exactly with the results derived from the work [3, 4], from our point of view it did not mean that one of this study was wrong, it only meant that pathologies, degrees of severity and gestational ages were different.

In any event, the superiority of the similarity entropy compared to the sample entropy reveals that the fluctuation around the FHR trend was the discriminating information. This result, that is completely novel in this field, is very encouraging for a discrimination purpose.

However, this work constitutes a first step in order to better understand the underlying processes leading to the foetal distress. The second step which is

foetus State	$RE_2(1,6)$	$RE_1(1,6)$
Healthy	41 %	51 %
Suffering	56 %	74 %

Table 1: Relative error of the similarity and sample entropies for different scales ($\alpha = 1$, $\alpha' = 6$).

in progress will consist to test our multi-scale descriptors to a larger corpus of recordings.

References

[1] Liston, R. and Sawchuck,D. and Young, D. (2007) 'Surveillance du bien-etre foetal: Directive consensus d'antepartum et intrapartum', *Journal of Obstetrics and Gynaecology Canada*, Vol. 29, pp. 1–58.

[2] Costa, M. and Goldberger, A.L. and Peng, C.-K. (2002) 'Multiscale entropy analysis of complex physiologic time series', *Phys. Rev. Lett.*, Vol. 89, pp. 068 102.

[3] Ferrario, M. and Signorini, M.G. and Magenes, G. and Cerutti, S. (2006) 'Comparison of Entropy-Based Regularity Estimators: Application to the foetal Heart Rate Signal for the Identification of foetal Distress', *IEEE Trans. BME*, Vol. 53, pp 119-125.

[4] Ferrario, M. and Signorini, M.G. and Magenes, G. (2009) 'Complexity analysis of the foetal heart rate variability: early identification of severe intrauterine growth-restricted foetuses', *Med. Eng. Comput.*, Vol. 47, pp 911-919, DOI 10.1007/s11517-009-0502-8.

[5] Xie, H.-B. and He, W.-X. and Liu, H. (2008) 'Measuring time series regularity using nonlinear similarity-based sample entropy', *Physics Letters A*, Vol. 372, pp. 7140-7146.

[6] Voicu, I. and Girault, J.-M. and Menigot, S.(2012) 'Improved estimation of the foetal heart rate using directional Doppler signal and YIN', *IRBM*, Vol. 33, pp. 262-270.

[7] Voicu, I. and Girault, J.-M. (2012) 'Multi-scale Similarity Entropy as a New Descriptor to Differentiate Healty to Suffering foetus', *International Conference on Complex Systems*, 5-6 Nov. Agadir-Morocco, 10.1109/ICoCS.2012.6458605.